

Haralampos Hatzikirou¹

Statistical mechanics of cell decision-making: the cell migration force distribution

¹ Braunschweig Integrated Centre of Systems Biology and Helmholtz Center for Infection Research, Inhoffenstr 7, 38124 Braunschweig, Germany, E-mail: haralampos.hatzikirou@helmholtz-hzi.de

Abstract:

Cell decision-making is the cellular process of responding to microenvironmental cues. This can be regarded as the regulation of cell's intrinsic variables to extrinsic stimuli. Currently, little is known about the principles dictating cell decision-making. Regarding cells as Bayesian decision-makers under energetic constraints, I postulate the principle of least microenvironmental uncertainty principle (LEUP). This is translated into a free-energy principle and I develop a statistical mechanics theory for cell decision-making. I exhibit the potential of LEUP in the case of cell migration. In particular, I calculate the dependence of cell locomotion force on the steady state distribution of adhesion receptors. Finally, the associated migration velocity allows for the reproduction of the cell anomalous diffusion, as observed in cell culture experiments.

DOI: 10.1515/jmbm-2018-0001

1 Introduction

Cell decisions are responses of the cell's intrinsic mechanisms to extrinsic signals [1]. Such mechanisms are typically termed signal transduction pathways. Typically, signal transduction occurs when an extracellular signaling molecule binds to a certain receptor creating a complex. In turn, this complex of molecules triggers a biochemical chain of events inside the cell [2]. These biochemical events can influence the epigenetic or even the genetic state of the cell. Then the new genomic state of the cell will be reflected in cell's transcriptome and will be translated into proteins that generate a phenotypic response. In general, cells encode extrinsic information into new intrinsic states and in turn decode the latter into phenotypic responses.

Concerning cell decision-making there are two fundamental issues: (i) typically there is a great deal of uncertainty in the intracellular dynamics involved in cell phenotypic decisions, and (ii) although there is a lot of work focused on single cell decision-making [1], in the context of multicellular systems, little is known. In particular, there are two schools of thought to address the single cell decision-making.

The first school thought is a *mechanism-driven* approach which follows the idea of epigenetic landscape reconstruction [3]. This dates back to Waddington's epigenetic landscape where cell phenotypes trickle down along a bifurcating landscape of "valleys" and "hills" (see Figure 1). Reconstructing such landscapes requires very good knowledge of the underlying genetic, epigenetic or transcriptional dynamics and provides elegant mathematical formulations that allow for temporal predictions. Obviously, this approach suffers from the issue (i), i.e. we do not know all possible intracellular players and dynamics involved in phenotypic decision.

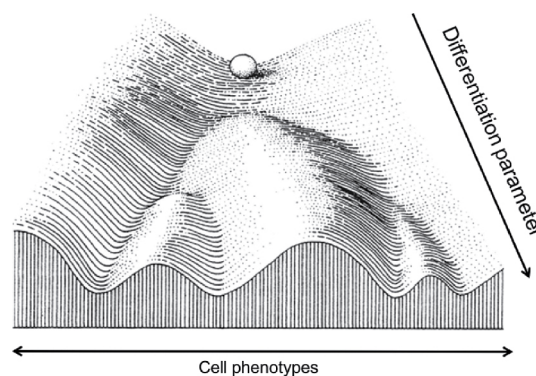


Figure 1: Sketch of communication between cell and its microenvironment.

Haralampos Hatzikirou is the corresponding author.
©2018 Walter de Gruyter GmbH, Berlin/Boston.

The second is a *data-driven* approach involving ideas from statistical inference [4]. In principle, cell decisions are treated in a mechanism-blind way, as an input-output system, without requiring the knowledge of the intracellular components involved in phenotypic regulation. The intracellular dynamics are treated as a black box. The main goal is to optimize quantities such as mutual information and infer the corresponding phenotypic responses. The main problem with this approach is that ignores dynamics, i.e. does not allow for temporal predictions or analysis of the cells phenotypic dynamics.

Given the above problems, my goal is to identify a variational principle for cell decision-making that fulfills the following: (i) keeps the elegance and predictability of a dynamic description, and (ii) circumvents the details of intracellular dynamics as statistical inference. Here, I propose the least microenvironmental uncertainty principle (LEUP) which I develop further in a statistical mechanics theory for cell decision-making. Finally, I apply the LEUP theory into the cell migration problem and I calculate the corresponding force distribution and migration velocity.

2 The Bayesian cell decision-maker

Cells typically collect information from their microenvironment and encode it into tangible actions, i.e. phenotypic decisions. The latter involves the rearrangement of their intrinsic variables $x \in X \subset \mathbb{R}^n$, representing genes, RNA molecules, translational proteins, metabolites, membrane receptors, etc. Microenvironmental or extrinsic variables are denoted by $y \in Y \subset \mathbb{R}^n$ being: ligand, chemicals, nutrients, cellular density or stress fields. To formalize this encoding process, we consider Bayesian cell decision-making:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}. \tag{1}$$

In particular, the extrinsic variable distribution $p(y/x)$ represents the likelihood of microenvironmental information being collected by the cell. In turn, it is combined with the prior probability distribution function (pdf) of intrinsic variables $p(x)$ resulting into the posterior intrinsic variable distribution $p(x/y)$. The latter represents the decision of the cell over the available information.

The collection of microenvironmental data, i.e. assessing the likelihood $p(y/x)$, is energetically expensive. For instance, as shown in Figure 2, sampling the cellular microenvironment involves the recruitment/degradation of different receptors (intrinsic state) and the corresponding binding to diffusible ligands (extrinsic state). As in any physical system, cells try to minimize energetic costs unless otherwise required. Therefore, I postulate that an energy-efficient cell would require the construction of the least informed/optimal prior.¹ This idea is rather straight forward in the case of differentiated cells, as heart cells, for instance, sufficiently know their heart microenvironment and they would fail living in another one, e.g. skin.

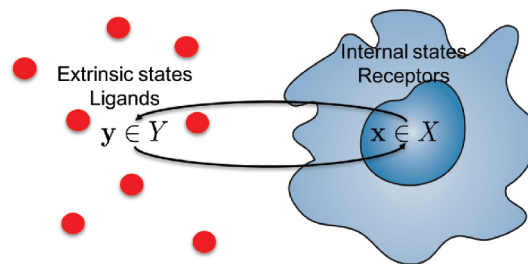


Figure 2: Sketch of communication between cell and its microenvironment.

3 A free-energy principle: the least microenvironmental uncertainty prior

The mentioned arguments could be translated in a statistical mechanics formulation. My central idea is that the cell decides on its phenotype by virtue of minimally sampling its microenvironment. As stated here, my goal is to optimize the cell's prior and at the same time minimize the total uncertainty of the system cell-microenvironment. This can be mathematically translated into finding the appropriate intrinsic state pdf $p(x)$ that minimizes the joint entropy of cell intrinsic and extrinsic variables

$$\min_{p(x)} S(X, Y), \tag{2}$$

where $S(\mathbf{X}, \mathbf{Y}) := -\int dx dy p(\mathbf{x}, \mathbf{y}) \ln p(\mathbf{x}, \mathbf{y}) = S(\mathbf{X}) + S(\mathbf{Y} | \mathbf{X})$.

The corresponding variational principle that follows reads:

$$\frac{\delta}{\delta p(\mathbf{X})} [S(\mathbf{X}, \mathbf{Y}) - \lambda (\int dx p(\mathbf{x}) - 1)] = 0, \quad (3)$$

where the $\frac{\delta}{\delta p(\mathbf{X})}$ denotes the functional derivative operator and λ a Lagrange multiplier for the normalization constraint. The solution of eq. (3) is the equilibrium pdf:

$$p(\mathbf{x}) = \frac{e^{-\beta S(\mathbf{Y}|\mathbf{X}=\mathbf{x})}}{Z}, \quad (4)$$

where $Z = \int dx e^{-\beta S(\mathbf{Y}|\mathbf{X}=\mathbf{x})}$ the corresponding normalization factor. The last equation is very interesting as it shows the equilibrium distribution of the cell's intrinsic states is the one that confers the least uncertainty of microenvironmental conditions. Please note that I introduced the parameter β as an "inverse temperature", that quantifies the compliance of a cell to LEUP. Please note that additional biological knowledge will be translated into Lagrange constraints of the free-energy minimization.

To illustrate the above result, let us assume that the distribution $p(\mathbf{y} | \mathbf{x})$ follows the normal distribution $\mathcal{N}(\mu_y(\mathbf{x}), \sigma_y^2(\mathbf{x}))$. Then the mutual entropy reads

$$S(\mathbf{Y} | \mathbf{X} = \mathbf{x}) = \ln (2\pi\sigma_y^2(\mathbf{x}))^{1/2},$$

and the eq. (4), respectively, becomes:

$$p(\mathbf{x}) = \frac{\sigma_y^{-\beta}(\mathbf{x})}{\int dx \sigma_y^{-\beta}(\mathbf{x})}. \quad (5)$$

This last equation is exactly the normalized square root of Fisher's information for the second moment, i.e. $\sqrt{\mathcal{F}(\sigma_y(\mathbf{x}))} \propto 1/\sigma_y(\mathbf{x})$ [5]. Given the latter, our equilibrium distribution (5) coincides with a normalized Jeffrey's prior distribution, for $\beta = 1$. Jeffrey's prior represents the definition of non-informative Bayesian prior in classical statistics [6].

Substituting the equilibrium pdf (4) into the cell intrinsic state entropy, we obtain:

$$S(\mathbf{X}) = -\int dx p(\mathbf{x}) \ln \frac{e^{-\beta S(\mathbf{Y}|\mathbf{X}=\mathbf{x})}}{Z} = -\beta S(\mathbf{Y}|\mathbf{X}) + \ln Z. \quad (6)$$

The above is a fundamental thermodynamic identity of LEUP where the first term denotes the system's total phenotypic internal energy $U = -(S(\mathbf{Y}|\mathbf{X} = \mathbf{x}))_{p(\mathbf{x})}$ which is the average LEUP sensed by a cell. The last term represents a phenotypic free energy $F = -\beta^{-1} \ln Z$.

4 Cell migration force probability distribution against integrin receptors

To illustrate the implications of the LEUP in a single cell-microenvironment system, we assume the problem of single cell migration (see Figure 3). Cell membrane-expressed integrin receptors bind to the corresponding extra-cellular matrix (ECM) ligands. The formation of stable focal adhesion points (FA), i.e. bond clusters, is required to generate locomotion force. Cell migration starts when the force at the front of the cell is larger than the force at the rear, inducing the rupture of the rear cell bonds. Integrin receptors play the role of cell's intrinsic state and the density of ECM ligands the corresponding microenvironmental/extrinsic variable [7]. Here, I try to calculate the locomotion force distribution via the surface-expressing receptor distribution.

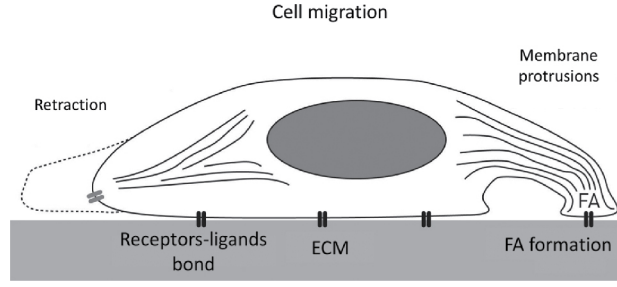


Figure 3: Simplified description of cell migration processes.

I employ the simplest FA dynamics model of receptor-ligand concentrations denoted by $x(t)$ and $y(t)$, respectively, and $c(t)$ corresponding to bond density:

$$\frac{dc}{dt} = k_+ y(x - c) - k_- c, \quad (7)$$

where k_+ the association and k_- the disassociation rate. Then, I calculate the steady state value of bond concentration:

$$c_{eq} = \frac{xy}{K_D + y}, \quad (8)$$

where the parameter $K_D = k_-/k_+$. It is well-known that the disassociation depends on the adhesion force F as $k_- = e^{F/c}$ and in the steady state one can prove (for details see in Schwartz and Safran [7]) that:

$$F_{eq} = p \log(k_+/e)x, \quad (9)$$

where the function $p \log$ is the solution of the equation $xe^x = a$. Therefore, the equilibrium force pdf is proportional to the steady state distribution of the receptors.

At this point, to calculate the equilibrium probability of receptors $p(x)$, I need to estimate the bond fluctuations of the system. The fluctuations estimated at equilibrium read:

$$\delta c_{eq} = \frac{dc_{eq}}{dy} \delta y = \frac{xK_D}{(K_D + y)^2} \delta y. \quad (10)$$

Using the Fourier transform $\delta \hat{c}_{eq}(\omega) = \int_0^T dt e^{-i\omega t} \delta c_{eq}(t)$, we can state that the standard deviation of c_{eq} is related to the power spectral density $|\delta \hat{c}_{eq}(\omega)|^2$ as

$$\langle \delta c_{eq}^2 \rangle = \frac{1}{2\pi} \int_{\mathbb{R}} |\delta \hat{c}_{eq}(\omega)|^2 d\omega.$$

The next goal is to calculate the steady state variation of complex formation. In this respect, we use the master equation as described in [8]:

$$\frac{dP_c}{dt} = k_+ y(x - (c - 1))P_{c-1} + k_- (c + 1)P_{c+1} \quad (11)$$

$$- [k_+ y(x - (c - 1)) + k_- (c + 1)] P_c, \quad (12)$$

where $P_c(t)$ the probability distribution of complexes and $c = 1, 2, \dots, x-1$. The resulting probability generating function reads at the steady state:

$$G_{eq}(s) = \left(\frac{s + K_D y}{1 + K_D y} \right)^x. \quad (13)$$

Then the complex distribution coincides with the Binomial one with parameter $(1 + K_D y)^{-1}$. The variance is simply given by the following formula:

$$\langle \delta c_{eq}^2 \rangle = \frac{K_D x y}{(K_D + y)^2}. \quad (14)$$

Combining equations (10) and (14), we can obtain the steady state fluctuations of the ligand density:

$$\sqrt{\langle \delta y(x)^2 \rangle} = (K_D + y) \sqrt{\frac{y}{K_D x}}. \quad (15)$$

Finally, assuming ligands for a given receptor concentration follow a normal distribution and a “inverse temperature” β , the equilibrium pdf of receptor is provided by eq. (5):

$$p(x) = \frac{\langle \delta y(x)^2 \rangle^{-\beta/2}}{\int dx \langle \delta y(x)^2 \rangle^{-\beta/2}} \propto x^{\beta/2}, \quad (16)$$

which results into power law distributions for different values of β . Assuming normality of the ligand y distribution and c_{eq} (the binomial distribution converges to the normal one for large numbers of receptors), then from the eq. (8) and for $K_D \gg 1$, we have $x \propto C_{eq}/y$. One can easily prove that the ratio of two Gaussian distributions results into fat-tailed Cauchy distribution which scales the data, i.e. $1/x^2$. LEUP would fit the corresponding data for $\beta = -4$.

This result states that the cell locomotion force pdf is proportional to the power law distribution of the adhesion receptors

$$p(F_{eq}) \propto x^{\beta/2}.$$

In particular, this implies that there is a finite probability for cells to exert high locomotion forces. In principle, the presence of power laws would support the long-standing hypothesis that biology is poised at criticality [9].

It is very interesting to see the impact on the cell locomotion velocity, where using Newton’s law the cell velocity reads:

$$v(t) = (mT)^{-1} \int_0^T F_{eq}(t) dt \propto \langle F_{eq} \rangle_t, \quad (17)$$

where m the cell mass and T the total time. Assuming ergodicity in the force generation process, then the above temporal average is equal to the ensemble average, i.e. $\langle F_{eq} \rangle_t = \langle F_{eq} \rangle_F$. Thus, the cell velocity reads, using eqs. (9, 16):

$$v \propto \int dF_{eq} F_{eq} p(F_{eq}) \propto \int dx x^{\beta/2+1} \propto x^{\beta/2+2}. \quad (18)$$

In the same result we could arrive by assuming that cells operate in the overdamped limit.

Now assuming $x \propto t^a$, where $a \geq 0$, one recovers the superdiffusive cell migration dynamics as observed in cell motility experiments [10] for $\beta > -4$ (see Figure 4), as

$$v(t) \propto t^{a(\beta/2+2)}.$$

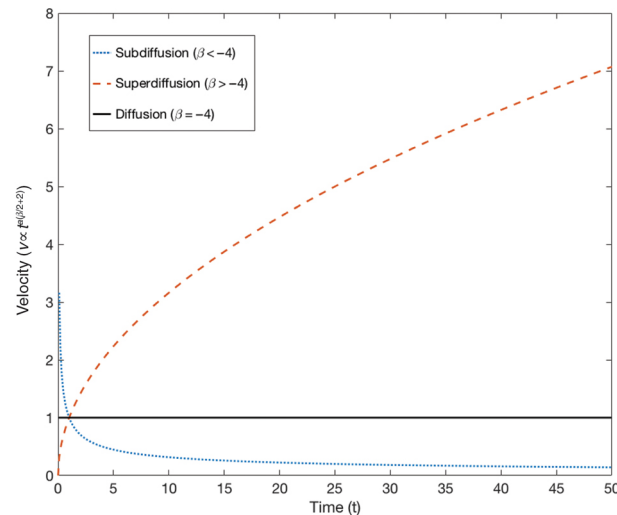


Figure 4: The LEUP compliance β controls the different regimes of cell diffusion.

In the low association limit $K_D \gg 1$, the LEUP compliance was estimated as $\beta = -4$. Then the average cell velocity is equal to a constant, i.e. the cell diffuses randomly. This is rather expected, as no stable focal adhesion points can be established, and the cell cannot establish a persistent migration direction. From the inverse point of view, random motion is not expected to decrease local microenvironmental uncertainty as LEUP dictates, therefore we could expect a negative β .

5 Discussion

In this article, I present a novel statistical theory for cell decision-making, the so-called LEUP. This is translated into a variational principle that minimized the joint entropy of cell intrinsic states and the corresponding microenvironmental states. In turn, I develop the statistical mechanics theory for equilibrium cell decision-making, where I derive thermodynamic-like relationships. In turn, I apply the theory in a simplified model of cell migration and I calculate the locomotion force distribution. Finally, the resulting force distribution recovers the observed single cell anomalous diffusion.

In the present paper, I develop only the equilibrium theory. However, the minimization of the LEUP free energy allows us to formulate cell decision-making dynamics. This can be in the form of Monte-Carlo simulations or Langevin-like stochastic differential equations. In a future work, I will elaborate on this direction.

One could note that this is not the first effort to use information theoretic arguments to understand cell fate determination. Similar approaches have been undertaken by other works, such as [11], [12], [13], [14]. However, in another paper I discuss the connection of all these approaches with LEUP [Hatzikirou H. *in preparation* 2018.].

Finally, LEUP is of particular interest as it allows for the inference of cellular intrinsic states/phenotypes by means of the local microenvironmental entropy or fluctuations. This allows for computation of the cellular states without the explicit understanding the underlying mechanisms. The sole knowledge of the extrinsic variable distribution is enough. Therefore, we could apply the LEUP to problems where phenotypic regulation is unclear or unknown. One direct application would be on the regulation of migration/proliferation plasticity or Go-or-Grow in the context of glioblastoma tumors [15], [16].

Notes

1 An informed prior is the one that requires the most sampling of new data. Here we are interested to least collection of data, i.e. the least informed prior.

References

- [1] Golding I. *Annu. Rev. Biophys.* 2011, 40, 63–80.
- [2] Purvis]E, Lahav G. *Cell* 2013, 152 (5), 945–956.

- [3] Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S. *Nature* 2008, 453 (7194), 544–547.
- [4] Cheong R, Rhee A, Wang C, Nemenman I, Levchenko A. *Science (NY)* 2011, 334 (6054), 354–358.
- [5] Frieden BR, Binder PM. *Physics from Fisher Information: A Unification* 2000, 68.
- [6] MacKay DJC. *Information Theory, Inference, and Learning Algorithms* David J.C. MacKay 2005, 100.
- [7] Schwarz US, Safran SA. *Rev. Mod. Phys.* 2013, 85 (3), 1327–1381.
- [8] Lauffenburger DA, Linderman JJ. *Receptors: Models for Binding, Trafficking, and Signaling*. Vol. 39. Oxford University Press, 1996:365. ISBN: 0-19-506466-6.
- [9] Mora T, Bialek W. *J. Stat. Phys.* 2011, 144 (2), 268–302.
- [10] Dieterich P, Klages R, Preuss R, Schwab A. *Science (NY)* 2008, 105 (2), 459–463.
- [11] Smith RCC, MacArthur BD. *Current Stem Cell Reports* 2017.
- [12] MacArthur BD, Lemischka IR. *Cell* 2013, 154 (3), 484–489.
- [13] Frieden BR, Gatenby RA. *PLoS ONE*. 2011, 6 (7), e22085–e22085. DOI: 10.1371/journal.pone.0022085.
- [14] Bialek W. *Biophysics: Searching for Principles*. Princeton University Press: Princeton, NJ, 2012.
- [15] Böttger K, Hatzikirou H, Voss-Böhme A, Cavalcanti-Adam EA, Herrero MA, Deutsch A. *PLoS Comput. Biol.* 2015, 11 (9), e1004366.
- [16] Hatzikirou H, Basanta D, Simon M, Schaller K, Deutsch A. *Mathematical Medicine and Biology*. 2010, 29 (1), 49–65. DOI: 10.1093/imamm-b/dqq011.